

SOLVING SOLUTION STRUCTURES OF PHYSIOLOGICALLY RELEVANT PROTEINS BY NMR SPECTROSCOPY

Neil E. MacKenzie, Paul R. Gooley, and Lori A. Hardaway

Department of Pharmaceutical Sciences, University of Arizona, Tucson, Arizona
85721

KEY WORDS: nuclear magnetic resonance, protein structure

INTRODUCTION

In the last decade or so, high resolution nuclear magnetic resonance (NMR) spectroscopy has become the technique of choice to study protein solution structure and dynamics at the molecular level. This is due, in part, to much improved instrumentation, but it is principally the consequence of the development of two-dimensional NMR methodologies. The latter were first proposed by Jeener in the early 1970s (56) and were vigorously applied to protein structural analysis some ten years later, initially by Wüthrich and his co-workers (134). Since then many groups have further developed NMR methodologies and applied them to the study of structure and dynamics of a wide array of proteins (see, for example, 9, 80, 129, 30, 57 for more recent reviews) and other macromolecules.

This chapter briefly summarizes the NMR protocol for the eventual realization of the three-dimensional solution structure of physiologically relevant proteins given the restraints of the NMR technique. Because of space limitations, this article is not exhaustive, and many fine studies are not mentioned in detail or at all (for example, 32). We have chosen to concentrate on families of related proteins on which detailed NMR structural analyses have been performed.

Protein Spectral Assignment Methodologies

The NMR techniques described below and many others that have been applied to various macromolecular structural elucidations have been the subject of a significant body of literature (see for example 134, 8, 38, 35, 61). There are, however, certain restrictions imposed on the physical properties of the protein to which these NMR techniques can be applied. For detailed structural analyses, a protein should have a correlation time (T_c) of approximately 10 ns. This factor is a function of the effective molecular weight (or hydrodynamic molecular weight at the protein concentration used) and, therefore, current structural studies are limited to those proteins with an upper M_r range of approximately 20 kd. The protein must be ipso facto nonaggregating at the practical NMR concentrations (1–2 mM) and should be stable at room temperature (or higher) for periods up to 48 hr.

Figure 1 is the one-dimensional ^1H NMR spectrum of polypeptide AII derived from pituitary bovine growth hormone (48, 46). Despite the relatively low molecular weight of this peptide, approximately (4 kd), the spectrum is still crowded and contains few resolved resonances. The assignment of a particular resonance to a specific ^1H nucleus in the polypeptide chain, the vital step in any structural analysis, is therefore almost impossible using one-dimensional NMR techniques alone.

A systematic two-dimensional NMR approach to the assignment of protein resonances has been developed by Wüthrich and co-workers (see 134 and references therein). This approach, with minor modifications (37), is essentially the standard for structural analyses of proteins (see conclusion for contemporary variations of this analytical approach). A recent review by Bax (9) discusses the protocols used to realize protein three-dimensional structures from two-dimensional NMR data and the following sections summarize these strategies.

Stage 1-Through Bond Correlations

The initial step in the ^1H resonance assignment process uses the phenomenon of scalar (spin-spin) interaction between neighboring hydrogens within two or three bonds of each other. The result is that resonances are split (J-coupled), with a magnitude that depends on the torsion angle between the C or N nuclei to which the hydrogens are attached. A series of two-dimensional NMR experiments have been developed that allows the identification of J-coupled hydrogens. The two-dimensional NMR experiment spreads the overlapping one-dimensional NMR spectrum of a protein into two orthogonal frequency dimensions and results in a dramatic improvement in resolution.

The COSY (correlated spectroscopy) experiment is one of the simplest, and detailed descriptions of the theory and application of this technique have been extensively published (for example 38, 5, 11, 49). A variation of this experi-

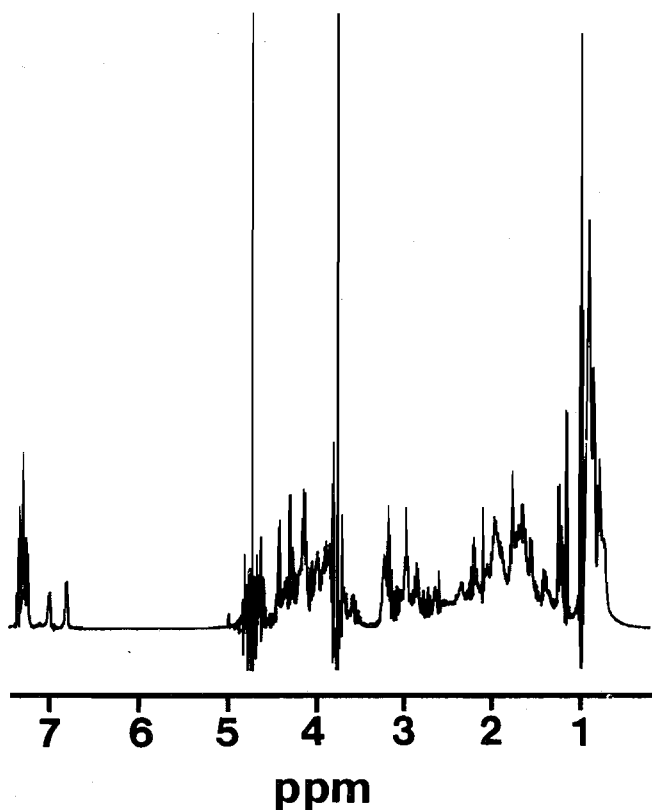


Figure 1 The one-dimensional 400 MHz ¹H NMR spectrum of polypeptide AII (2 mM) derived from pituitary bovine growth hormone. For a description of the acquisition parameters see Gooley et al (48).

ment involves multiple-quantum transitions that can select for particular amino acid types. The most commonly used is the double-quantum filtered (DQF) COSY (for example 108). Figure 2 is the DQF-COSY spectrum experiment of fragment AII and, as is the case in all COSY-type spectra, the off-diagonal crosspeak indicates connectivities between J-coupled hydrogens. A clearer example of quanta-filtering is shown in Figure 3A and B. Figure 3B is a section of the triple-quantum filtered (TQF) COSY (105, 114) spectrum of *Rhodobacter capsulatus* ferricytochrome *c*₂ (D. Zhao et al unpublished). Glycine has only two coupled protons when the amide hydrogen has been exchanged with deuterium, and therefore no triple-quantum transitions can be generated. Thus glycine C_αH hydrogens, which are present in the DQF-COSY spectrum (Figure 3A), can be selectively filtered in the TQF-COSY

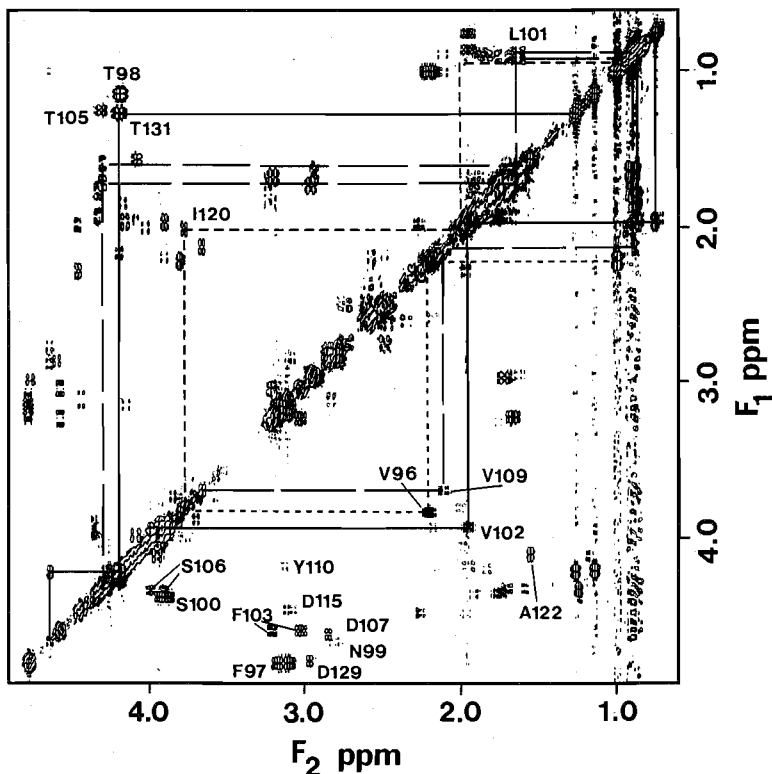


Figure 2 Contour plot of a region of a DQF-COSY spectrum of polypeptide AII (8 mM). The J-connectivities between the $C_{\alpha}H$ and $C_{\beta}H$ protons of the nine AMX (see Reference 134) spin systems (Phe-97, Asn-99, Ser-100, Phe-103, Ser-106, Asp-107, Tyr-110, Asp-115, and Asp-129) are indicated. In addition, the complete spin systems of the three Val (Val-96, Val-102, and Val-109), Leu-101, and Thr-131 are shown. The $C_{\beta}H-C_{\gamma}H$ cross-peak of Thr-98 and -105 and the $C_{\alpha}H-C_{\beta}H-CH_3$ J-connectivities of Ile-120 are also described. For a description of the acquisition parameters see Gooley et al (46).

experiment (Figure 3B). The obvious spectral clarification and reduction in overlap in the TQF-COSY spectrum can be seen and, despite the lower signal to noise ratio, this is a valuable tool for protein spectral assignment.

A third class of J-correlated experiments is referred to as total correlation spectroscopy (TOCSY or HOHAHA) which, in principle, allows the correlation of all protons within the spin-spin system that represents the individual amino acids. A detailed description of the use of this technique is given by Bax (10). After acquiring a series of J-correlated spectra, the initial goal in the total spectral assignment process is to subdivide the various spectra in terms of individual spin-systems for those amino acids that possess either a unique spin-spin signature or those that are indicative of a certain amino acid type (134, 60).

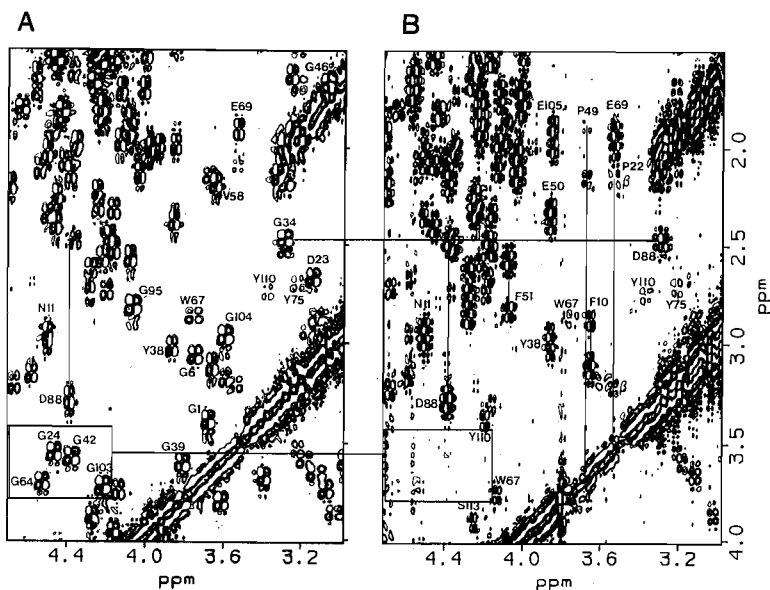


Figure 3 Contour plot of a region of a DQF-COSY (Figure 3A) spectrum and the same region of a TQF-COSY spectrum (Figure 3B) of *Rhodobacter capsulatus* ferricytochrome c_2 . Salient features of these spectra are the elimination of glycine $C_{\alpha}H$ connectivities in the TQF-COSY experiment as shown, in the boxed region. Note that the assignment problems caused by overlapping $C_{\alpha}H$ connectivities of Gly-34 and Asp-88 in the DQF-COSY experiment (Figure 3A) are resolved in the TQF-COSY spectrum (Figure 3B). D. Zhao et al (unpublished results) have given a description of the acquisition parameters.

Stage 2-Through Space/Sequential Correlations

In the NMR experiment, spatially proximal hydrogens mutually exchange their nuclear magnetization at a rate (k) that is inversely proportional to the sixth power of the distance (r) between them and directly proportional to T_c . These exchange rates can be measured and a set of interatomic distances can be obtained. This exchange phenomenon is called the nuclear Overhauser effect (NOE) (87). The application of the two-dimensional NOE experiment, NOESY, has been well documented (134 and references therein), and the NOE data obtained can be grouped into two general classes: short range and long range. The former, generally limited to amino acids that are less than five residues apart, is used to place an amino acid, or amino acid type previously characterized, in the J-correlated types of experiment as described above, into its sequential position in the polypeptide chain. The long range NOEs are used to characterize the global fold of the backbone of the protein (see below). For short mixing times (~ 60 ms), NOEs can be observed between hydrogens that are maximally 5 Å distant. Obviously

hydrogens that are more proximal than this will give rise to more intense NOEs. The r^{-6} dependence of the magnetization transfer rate implies that NOE data can be used to accurately determine interhydrogen distance; however, for reasons that are summarized in a following section, this is generally not the case.

NOE data can be used to qualitatively characterize secondary protein structural features. Thus in any secondary structural element, the distances and orientation between backbone hydrogens ($C_{\alpha}H$ and NH) and C_{β} hydrogens on adjacent amino acids are representative of that structural element. As exemplified in Figure 4, the α -helix is mainly characterized by intense sequential amino acid $N_{(i)}H$ to $N_{(i+1)}H$ interactions (the inter-NH distance in an α -helix is approximately 2.8 Å. Other secondary structural elements such as the parallel and antiparallel β -sheet, the 3_{10} -helix, various types of turns (I, II, IV, etc.), and omega-loop have their own characteristic NOE profiles. The

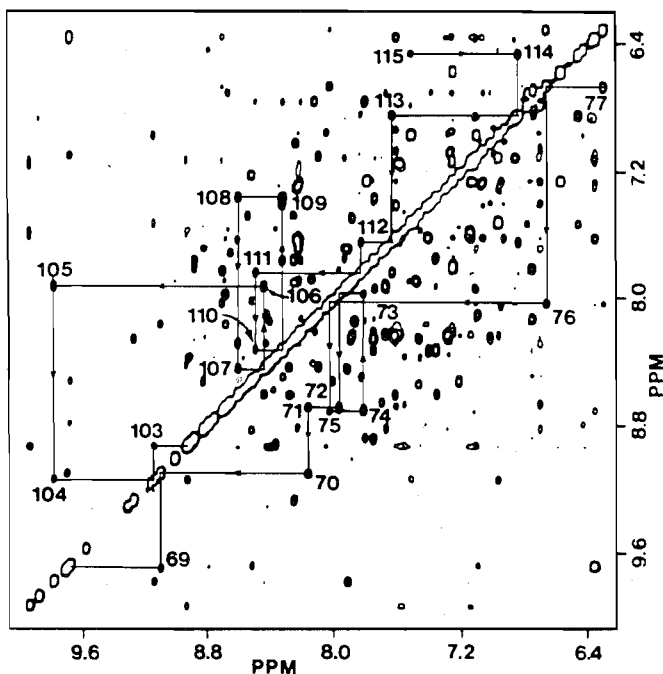


Figure 4 Section of a NOESY spectrum of *Rhodobacter capsulatus* ferredoxin I. The solid lines indicate $N_{(i)}H$ to $N_{(i+1)}H$ connectivities and describe a part of a long central helix from Glu-69 to Asp-78 in the lower half of the spectrum and the C-terminal helix Gly-103 to Lys-116 in the upper half. For a description of the acquisition parameters see Gooley et al (45).

majority of these profiles have been described by Wüthrich and co-authors (134, 124) and have been utilized extensively by many groups.

Additional parameters that more accurately define a secondary structural unit can be obtained by measurement of the J-coupling constants of hydrogens that are three bonds apart and by stereospecific assignments of certain methylene and methyl resonances. The magnitude of the former has a characteristic dependence on the dihedral (torsional) angle between the two hydrogens and the dependence can be approximated, in the case of polypeptides, by the Karplus equation (59). For example, because of the characteristic fixed orientation of the backbone atoms in typical secondary structural units, certain J values are representative of that unit. Thus the three-bond NH-C α H J value in an α -helix is approximately 4 Hz, while in an antiparallel β -sheet, this value is nearer 9 Hz (134).

There are certain drawbacks in using NOE intensities to define structure (see also next section). One such difficulty is encountered when one of the interacting partners giving rise to an NOE is a methylene or methyl group. For example, with methylenes it is difficult to ascertain which of the hydrogens actually participates in the interaction. The practice in any structural computation is to assume that a center, which resides in a geometric average position between the two hydrogens, is the interactive species. In this case, however, a distance error of approximately 1 Å is inherent in the assumption (133, 31, 53). It is easy to envisage that a similar handling of the methyl groups of a valine residue would be required. Recently, however, it was demonstrated that for valyl methyl groups and β -methylene sites (136, 55) stereospecific assignments might be possible and thus the resulting structural determination could be made more accurately.

Stage 3-Protein Tertiary Structure Determination

Once a complete or nearly complete sequential spectral assignment has been accomplished and the position and type of secondary structural units characterized, long-range NOE information is used to juxtapose these units in the global arrangement of the protein.

At this point, the two main drawbacks to the use of the NOE intensities as determinants of interatomic distances, alluded to previously, require further discussion.

The first is that k , the rate of magnetization transfer between two hydrogens (see above) has a positive value for those molecules with small T_c (e.g. 10^{-11} sec, typical of di and tripeptides) and a negative value for those molecules of $T_c > 10^{-9}$ sec, typical of proteins studied by NMR spectroscopy. A protein cannot, under physiological conditions at any one time, be described by a single T_c , and thus any single measurement of k in a NOESY experiment will

not accurately reflect the true value of the interatomic distance between all interacting hydrogens. For example, long amino acid side chains can be subject to rapid local motions (e.g. rotational), and these moieties will tend to have a relatively reduced k when measured optimally for the macromolecular T_c . Furthermore, interhydrogen distances may be modulated by slow local motions, again causing localized variations in T_c .

The second problem in determining the true magnitude of the NOE interaction is that if two hydrogens are within 2–2.5 Å of each other, they will give rise to an intense NOE, resulting from direct magnetization transfer. However, once this transfer has been accomplished, it may be further, or indirectly, transferred. The net result is that this diffusion of magnetization may give an inaccurate picture of the spatial orientation of the hydrogens involved (see Borgias & James, 20, for a recent, more detailed discussion).

In practice, data acquisition parameters are chosen to minimize the above drawbacks, and the distances obtained from NOE data, used in the structural computations (see below), are input as a range of values. For example, intense, medium and weak NOEs are ascribed to distances of 1.8 to 2.7, 1.8 to 3.3, and 1.8 to 5 Å, respectively.

Notwithstanding, these NOE distance constraints are used in tandem with those of bond angles, bond lengths, chirality, van der Waal's radii, J values, and hydrogen-bonding as input to generate three-dimensional protein structures. Two approaches are generally used: distance geometry and restrained molecular dynamics. The relative merits of both have been recently discussed (70, 115, respectively) and further discussion is beyond the scope of this review. Suffice it to say that a global representation of a protein can be realized in this fashion. As a consequence of the range of distance boundaries input from NOE data, families of structures, and not a unique solution, are generated. The consistency of this approach is measured by the magnitude of the root mean square (RMS) deviation of the three-dimensional coordinates between the members of the family. Typical, acceptable RMS values are in the region of 1.7–2 Å, and examples are given below.

PROTEIN STRUCTURES

Toxins

Polypeptide neurotoxins have proved useful in the physiological study of the transmission of neural signals. The two classes of ion channels involved in the generation of electrical signals are the ligand-gated channels, such as the nicotinic acetylcholine receptor, and the voltage-sensitive ion channels, such as the Na^+ channel (28). Various small toxic proteins from a wide range of sources have been used to study these channels, and many of these toxins

have been characterized successfully by NMR spectroscopy. These proteins are quite amenable to NMR analysis because they are small (less than 80 residues), globular, (due to the presence of one or more disulfide bridges), and stable to a wide range of solution conditions, including pH and temperature, with little or no requirement for buffer ions. For example, the bee toxin, apamin, has been studied for over a decade, and its solution structure and dynamics have been well characterized by both one and two-dimensional NMR techniques (24, 34, 101, 127). Other toxins that have been studied include snake venoms: α -bungarotoxin (7), α -neurotoxin (71), and CTXIIa and CTXIIb (94); scorpion toxins: I₅A from *Buthus eupeus* (4), variant-3 neurotoxin from *Centruroides sculpturatus* Ewing (85), and charybdotoxin from *Leiurus quinquestratus hebracus* (19); α -conotoxin from marine snails (99, 66); sea anemone toxins: AP-A (76), ATXIa (132), HpII (128), HpIII (100), and ShI (41). The extensive studies of the sea anemone toxins permit detailed comparisons and are discussed below. Several pertinent issues, such as purity of sample and correctness of amino acid sequence, have been encountered in the course of studying several snake toxins, and these will be summarized. The cardiotoxins of the venom of the cobra *Naja mossambica mossambica* have been studied for some time (116, 54). In the process of obtaining complete NMR spectral assignments, it was realized that the sample was a mixture of toxins (93). Thus only recently have the complete assignments for two of these proteins, CTXIIa and CTXIIb, been obtained (94), and as a result, only the secondary structure has been fully characterized. Both proteins possess a short double-stranded antiparallel β -sheet near the N-terminal and a large segment of triple-stranded antiparallel β -sheet. The long neurotoxin α -bungarotoxin from banded krait *Bungarus multicinctus* has been extensively studied by NMR spectroscopy (6, 7, 69). From the sequential assignments, four errors in the order of the amino acid sequence were detected, which were subsequently confirmed by chemical sequencing. More importantly, significant differences were observed between the solution and crystal (by X-ray analysis) structures, which included a larger β -sheet and a difference in the orientation of the conserved Trp. These differences are believed to be real and reflect the supposition that there is more than one stable solution conformation of α -bungarotoxin. This flexibility may be important for the recognition at the acetylcholine receptor.

Sea Anemone Toxins

A series of small (approximately 50 residue) polypeptide toxins isolated from sea anemones show differing degrees of cardio and neurotoxicity. They bind to the neuronal voltage-gated Na⁺ channel, thereby slowing down channel inactivation and prolonging the action potential. On the basis of amino acid sequence, immunological cross-reactivity, and competition studies for bind-

ing to the Na⁺ channel, two classes of sea anemone toxins can be distinguished: the actiniid proteins are classified type 1 and include ATXI, ATXIa, ATXII from *Anemonia sulcata*, and AT-A from *Anthopleura xanthogrammica*; the stichodactylid proteins are classified type 2 and include ShI from *Stichodactyla helianthus* and HpI and HpII from *Heteractis paumotensis*. Near complete ¹H assignments have been published for AP-A (47, 76), ATXIa (132) HpII, (128), HpIII (100), and ShI (41, 88).

Solution structures for ATXIa (131) and ShI (40) are well-defined, and preliminary structures for AP-A have been obtained (119). The structural characterization of AP-A has been hampered by apparent *cis-trans* isomerism of the peptide bond of Pro-41 (43). All four proteins contain a core of anti-parallel β -sheet encompassing residues 1-3, 19-24, 29-36, and 40-47, connected by two well-defined loops, and a third poorly defined loop that encompasses residues 7-16 (using ShI numbering, see Figure 5). The β -sheet is distorted by the three common disulfide bridges. Surprisingly, the tertiary structure of ShI (a type 2 toxin, Figure 5) is more like ATXIa (type 1) than HpII (type 2). The greatest difference between ATXIa and ShI is that the loop 34 to 39 packs more closely to the β -sheet in ShI than in ATXIa, a consequence of this loop having several tight turns in ShI and being more extended in ATXIa. The characterization of the solution structure of BDS-I, another small peptide from *Anemonia sulcata*, shows that this protein contains a β -sheet core and disulfide bridge network similar to the type 1 and 2 toxins (36). BDS-I shows poor sequence homology to the type 1 and type 2 neurotoxins and, in addition, it is an antiviral and antihypertensive agent, but is neither neuro or cardiotoxic, which suggests that the basic fold of these proteins may not be important for toxicity.

Rather than the basic fold being important for the activity of these toxins, perhaps other subtle molecular interactions, such as hydrogen bond stability and other dynamic phenomena, are important for function, and a number of studies probing these phenomena have been conducted. For example, the conformational pH dependence of ShI differs significantly to AP-A, ATXI, and ATXII (88, 44). ShI unfolds at alkaline pH, whereas the latter polypeptides are stable up to pH 13. pH titrations of AP-A, ATX I, and II suggest that a structurally important carboxylate, possibly Asp-9, which is not a part of the β -sheet, titrates with a pKa near 2 for AP-A and ATX II and a pKa near 3 for ATX I, where ATX I is the least toxic (44). A carboxylate with a pKa of 2.8 is also observed in ShI, although this titration is not assigned to Glu-8, the equivalent residue to Asp-9 of the actiniid proteins. In addition, the thermal stability of ShI is lower than AP-A, ATX I, and ATX II (88, 44), which together with NH exchange studies show ShI to be less stable than AP-A and ATXI (120). Further investigations, including single residue replacements, either synthetically or by site-directed mutagenesis, may further an understanding of the mode of action of these proteins (103).

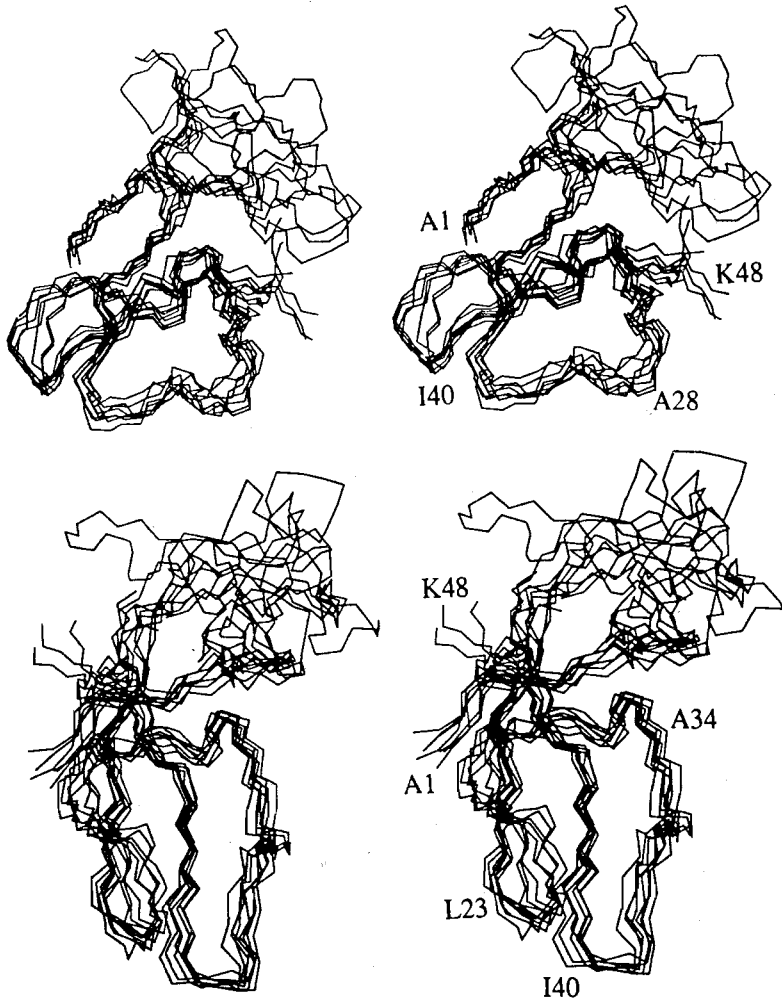


Figure 5 Backbone atoms of a family of eight structures of neurotoxins from the sea anemone *Stichodactyla helianthus* (ShI). Pair-wise RMS displacement values for these atoms are 1.3 ± 0.2 Å when the poorly defined loop region (residues 7–16) is omitted from the calculation. The top and bottom halves of the figure show the structures viewed from two different angles (40). Left and right are stereo pairs.

DNA-Binding Proteins

Over the last few years an increasing number of studies have shown that the binding region of proteins believed to be involved in the recognition of DNA sequences can generally be classified into a number of structural motifs. Three well-described structures are the helix-turn-helix, the zinc finger, and

the leucine zipper, all of which are discussed below. It is important to note that many DNA-binding proteins do not fit into these categories, and therefore examples of other DNA-binding proteins are also described.

ZINC FINGERS The zinc finger motif has been identified as a nucleic acid-binding domain. These motifs have been further categorized into two classes, where the Zn ion in the classical zinc finger is coordinated by two Cys and two His residues, and in the second class the ligands are four Cys residues. Three-dimensional structures have been described for four synthetic peptides, each corresponding to a classical single zinc finger, but from different sources: the *Xenopus* protein Xfin-31 (74, 75), corresponding to the 31st zinc finger of this transcription factor; the yeast transcription activator ADR1 (65); a human enhancer-binding domain (HEBD) (90); and the mouse Kruppel-like gene mKr2 (26). With the exception of the latter protein, the samples were prepared with at least a stoichiometric amount of zinc. Not unexpectedly, all four proteins show structural similarities to each other, but there are also a number of differences. The Xfin-31 finger starts with two β strands arranged in a hairpin structure. This structure is not observed in the other zinc fingers, where an irregular, antiparallel β -sheet connected by an atypical turn is observed (see Figure 6). In each finger, a helix of varying length is observed: 12 residues in Xfin-31, 12 to 13 in ADR1; 10 in HEBD (residues 11–24, Figure 6); and 7 in mKr2. In each protein, except mKr2, the helix starts at the same residue. In mKr2 the helix appears to start approximately 4 residues later, which accounts for this helix being shorter and possibly a consequence of the protein samples being prepared without zinc. In the ADR1 and Xfin-31, but not the mKr2 and HEBD fingers, the helix includes both ligand His residues. The helix appears to be all α in the ADR1, mKr2, and HEBD zinc fingers (Figure 6). However, in Xfin-31 the helix starts as an α -helix and after the first His ligand, as a number of i to $i + 2$ connectivities are observed, finishes as 3^{10} -helix. NOE data for the Xfin-31, ADR1, and HEBD proteins are consistent with the N ϵ atom of the two His ligands forming the zinc ligand bonds.

The two zinc fingers within the segment Cys-440 to Arg-510 of the rat glucocorticoid receptor DNA-binding domain are class 2. Similarly, the DNA-binding domain of the estrogen receptor also consists of two class 2 zinc fingers. Three-dimensional structures of fragments containing the two fingers for both receptors have been calculated (52, 113). The structure of these fingers differs significantly from the classical fingers in that no helical regions were found within the fingers. For the glucocorticoid receptor fragment, a small piece of antiparallel β -sheet is found in the N-terminal region of the first finger, and a type I and type II turn encompasses the N-terminal Cys ligands of the second finger. α -Helices and several portions of extended conformation to the C-terminal side of each finger are observed such that a pattern finger-

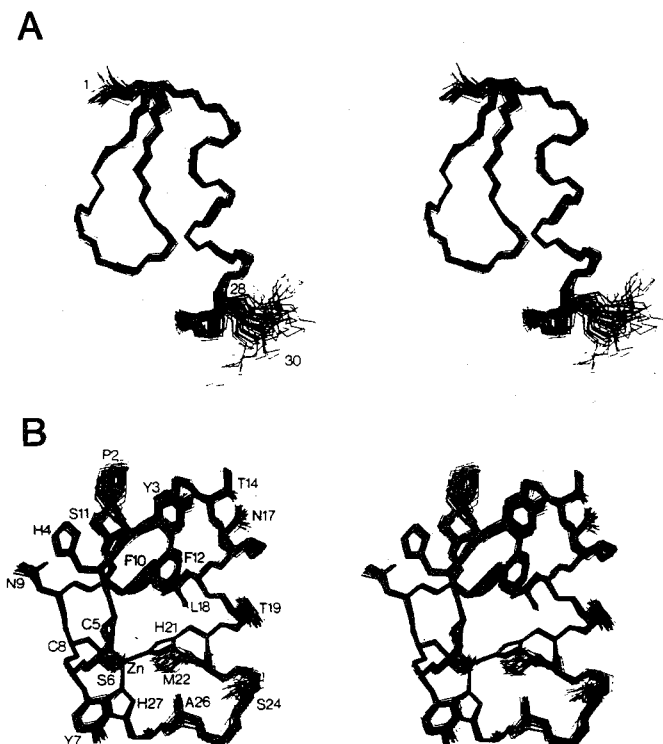


Figure 6 (A) Best fit superposition of the backbone atoms of a family of 40 structures of the 30-residue single zinc finger from a human enhancer-binding domain (HEBD). (B) Best fit superposition of the backbone and side-chain atoms of a family of 40 structures of the first 27 residues of HEBD. The C-terminal tripeptide is poorly defined (Figure 6A). The RMS displacement for all atoms in the absence of this region is $0.41 \pm 0.07 \text{ \AA}$ (90). Left and right are stereo pairs.

helix-extended is repeated. NOE data show a number of long range interactions between the two fingers and between the two helical domains. In three-dimensional structure, calculations show that the two helices pack perpendicular to each other with hydrophilic faces exposed to solvent, one finger contacts with the two helices, and the second finger extends out from the protein with the two zincs, 13 \AA apart. In this second finger there is an exposed region, Ala-477 to Asp-481, proposed to be involved in protein interactions between monomers of the glucocorticoid receptor and thus deemed important for cooperative binding. The structure calculations of the estrogen receptor DNA-binding domain show similar structures to those described for the glucocorticoid receptor. However, one region of difference appears to be a nine residue loop that may bind to the minor groove (113).

These studies of the two classes of zinc fingers show differences in structures and therefore the hypotheses of DNA-binding differ. In the study of the class 2 zinc fingers of the glucocorticoid receptor, the helix to the C-terminal end of the first finger was successfully modeled to bind to the DNA major groove (52). For the classical fingers, two models for DNA-binding have been proposed: (a) where successive fingers make structurally equivalent contacts with the major groove without crossing the minor groove (12), (b) where alternative fingers bind on one face of the major groove, so that successive minor grooves are crossed (39). In the three classical fingers studied, a number of positively charged and polar side chains are located on the solvent exposed side of the helical region and are speculated to be involved in DNA binding. Using data from hydroxyl radical footprinting and the structure of the zinc finger (75), a model supporting the second mode of binding has been proposed (29). Experimentally, the classical fingers that have been discussed above are single isolated fingers, and because the proteins from which these fingers are derived contain multiple fingers, it is important to address how the structure of individual fingers relate to each other. The transcription factor SW15 contains three fingers near its C-terminus. Studies on constructs of the first two fingers show that similar NOEs are observed in either the one or two finger constructs, which suggests that the individual fingers are modular (86). In the course of determining the structure of HEBD (90) zinc finger, the authors noted that, despite a lack of sequence homology, this zinc finger showed striking structural homology to residues 23–44 of the proteinase inhibitor ovomucoid third domain from Japanese quail. This observation suggests that the classical zinc finger structure is not new, but may be found in a variety of proteins, thus supporting the notion that tertiary folding motifs can be limited to a number of classes.

HELIX-LOOP-HELIX The structures of a number of DNA-binding proteins, for example, the λ (95), *cro* (2), and *trp* (112) repressors, solved by X-ray crystallography, show a protrusion of a short α -helix, a turn followed by a second α -helix. This structural motif is involved in the recognition of the specific base sequence of the operator where the second helix lies partly within the major groove in contact with the specific bases and is therefore termed the recognition helix. The tertiary structures of this class of proteins are difficult to solve by conventional two-dimensional NMR methods because of their molecular weight. However, sequential assignments of the *trp* repressor, a symmetric dimer of 25 K, have been determined (3). For this protein, assignments were obtained by selectively deuterating various amino acid residues and observing their effect on the intensity of NOE crosspeaks. In other cases, fragments containing the DNA-binding domain can be prepared and studied, a method particularly useful in obtaining structural data for repressor proteins that have not been crystallized. For example, the three-

dimensional structure of the lac repressor headpiece (136, 58, 122) and the Antennapedia homeodomain from *Drosophila melanogaster* (107, 13) have been determined, each showing a helix-turn-helix motif. Comparison of the latter protein to other repressors shows that the recognition helix is slightly longer than the analogous helix of these other proteins.

An NOE analysis of a complex of the lac repressor headpiece with a 14 base-pair fragment of the lac operator showed that, as predicted, the recognition helix binds in the major groove, but the orientation of this helix is 180° different to that of other repressor-operator complexes (17). A similar study, but on a 2:1 complex of lac operator headpiece with a 22 base-pair symmetric operator, showed that each headpiece bound the operator in the same way as the previous study with the 14 base-pair operator (72). These studies suggest that the helix-turn-helix proteins may be further subdivided into how the recognition helix binds to the specific operator sequence.

LEUCINE-ZIPPERS AND OTHER DNA-BINDING PROTEINS A number of other DNA-binding proteins do not contain helix-turn-helix or zinc finger motifs. A recently proposed motif, the leucine zipper, has been observed in several proteins, e.g. the transcription factors GCN4, *fos* and *jun* (91, 92), and a histone DNA-binding protein (117). The leucine zipper is characterized by a helical region of 30 residues with a periodic repeat of Leu every 7 residues (73). Relatively little structure knowledge has been obtained for these proteins, although synthetic peptides corresponding to the zippers of GCN4, *jun* and *fos* form helical dimers as determined by circular dichroism (91, 92). Two-dimensional NMR analysis of a synthetic peptide corresponding to the zipper of the yeast transcriptional activator GCN4 shows that the peptide forms a symmetric dimer, consistent with a coiled-coil, and is α -helical for almost the entire peptide, e.g. 32 of 33 residues (89). It is important to note that the DNA-binding domain of GCN4 consists of two subdomains, a basic region, which is the DNA recognition site, and the leucine zipper. Structure calculations of a 60-residue peptide encompassing these regions showed NOEs, including medium range NOEs (i to $i + 3$), characteristic of helix for the leucine zipper region (111). In contrast, few interresidue NOEs were observed for the basic region, which suggests that it is flexible, although the authors reported that, at longer mixing times, NOEs characteristic of helix could be observed. Circular dichroism and NMR studies of a 56-residue fragment (225–281) of GCN4, which includes both subdomains, suggest that the basic region is unstructured in the absence of DNA and helical in its presence (125). These data suggest that the flexibility of the recognition domain is important for DNA binding.

The Arc and Mnt repressors of the bacteriophage P22 show little sequence homology to the helix-turn-helix proteins. From the sequence-specific assignments of the Arc repressor, which is a 53-residue dimer, and a thermostable

proline to leucine-8 mutant, the secondary structure has been determined as extended-helix-turn-helix for both proteins, where the extended chain and the two helices are amphipathic (22, 135). NOE data suggest that the extended strand may form a β -sheet, but it has not been determined if the β -sheet is inter- or intra-monomer. The N-terminal region is implicated in operator recognition, thus suggesting that the DNA-binding motif is β -sheet and that the Arc and probably the homologous Mnt repressor belong to another class of DNA-binding proteins.

The DNA-binding domain of the transcription factor GAL4 was previously thought to contain two zinc finger motifs and thus binding two zinc ions each with Cys2 and His2 residues. Recent ^{113}Cd NMR studies on apoprotein fragments of GAL4 reconstituted with ^{113}Cd show two ^{113}Cd -binding sites (96–98). However, on the basis of ^1H - ^{113}Cd -coupling constants, a total of six Cys residues were identified as metal ligands, of which two form bridging ligands to both ^{113}Cd ions. Therefore, this DNA-binding domain is a binuclear metal ion cluster rather than two zinc fingers. These data are in agreement with EXAFS studies (106). A 43-residue fragment of GAL4 (residues 7–49) was found to contain two zincs and bind DNA (42). NOE analysis of this fragment showed connectivities compatible with helices, although the position of these helices differs from that predicted for a zinc finger.

HORMONES

Insulin

Insulin is involved in a myriad of biochemical pathways, particularly the mobilization of glucose. ^1H NMR spectroscopy of this protein has been hindered by its tendency to self-associate into multiple exchanging forms. Recently Kline & Justice (64) assigned the ^1H NMR spectra of human insulin, wherein an organic solvent is used to decrease self-association (126). In acetonitrile at low pH, the conformation of human insulin was found to be consistent with that determined by X-ray crystallography (Figure 7) (1) except for discrepancies at the C-terminus of the B chain. X-ray studies have characterized an extended structure for this region, which was not detected in the NMR study (Figure 7). This discrepancy may reflect real differences between solution and crystal structures. Under physiological conditions (high dilution, pH 8.0 and 9.3) Roy and co-workers (110) found a conformational difference between the monomer and aggregated forms.

The discovery of variant insulin in patients with diabetes (83 and references within) has stimulated research in the study of insulin analogues, including Asp B9, Leu B25, des-pentapeptide, (des-B26-B30), and mini-proinsulin (an analogue containing a peptide bond between residues B29 and A1) (110, 16, 126, 21, 82).

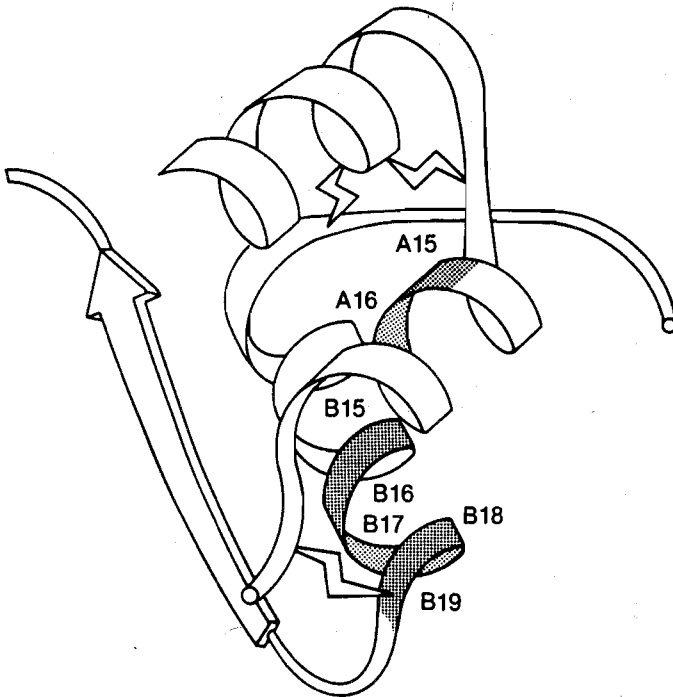


Figure 7 Ribbon diagram of the three-dimensional structure of insulin (based on the X-ray coordinates). The shaded helical regions in the A and B chains indicate the most stable regions of this protein as assessed by amide hydrogen \rightarrow deuterium exchange rates (64).

Growth Factors

Excellent articles covering structure-function relationships of growth factors and their receptors (81) and of epidermal growth factor (EGF) and transforming growth factor alpha (TGF-alpha) have been recently published (25). In brief, urogastrone [human epidermal growth factor (hEGF)], a 53-amino acid peptide with three disulfide bridges, acts primarily to promote cell growth and differentiation. hEGF has been studied by ^1H NMR spectroscopy (27, 33). The dominant structural feature was found to be a pair of antiparallel β -sheets. Human transforming growth factor (hTGF-alpha) is composed of 50 amino acids and is related to the EGF family in both sequence and activity. It is suggested that the "mitten model" structure of TGF-alpha as proposed by Kohda, (68), previously used to describe mouse EGF (67), is the functionally relevant structural motif.

Other Hormones

Other physiologically relevant hormones such as glucagon, secretin, and gastrin have also been studied. Due to their small size, less than 40 residues,

and lack of disulfide bonds, it is predicted that in aqueous medium they may not adopt a particular conformation.

This is typified in the case of glucagon, a 29 amino peptide that promotes glycogenolysis. Boesch and co-workers (18) found that glucagon, in dilute aqueous solution, is essentially disordered with perhaps the exception of residues 22–25. However, in the presence of perdeuterated dodecylphosphocholine (DDP) micelles, certain conformations of glucagon can be stabilized (23, 130). These are an irregular strand (residues 5–10) followed by a helix (residues 10–13), a half turn (residues 14–16), and a helix in residues 17–29 (51).

Secretin is a 27-residue peptide that mediates the secretion of alkaline pancreatic juices. It is a member of the glucagon superfamily and has 52% amino acid sequence homology with glucagon. Again, in aqueous solution the conformation of this peptide as analyzed by CD is predominantly disordered (14, 15, 109). The addition of trifluoroethanol (TFE) up to a concentration of 35% increases helical content of secretin. At 40% TFE, secretin shows remarkable structural similarities to glucagon in DDP micelles, i.e. an irregular strand is seen in residues 2–6, a helix (residues 7–13) a half turn (residues 14–16), and another helix spanning residues 17–25 (51).

Gastrin exists as 14, 17, and 34 residue homologues, which all have a common tetrapeptide C-terminal sequence. This common tetrapeptide exhibits all the effects of the larger homologues, although biological potency has been found to increase as chain length increases towards the N-terminus up to the nonapeptide (121). Although there has been disagreement about the conformation of gastrin, especially at the C-terminus, the N-terminus has been shown to fold into a helix, which may be important for receptor recognition (118).

Similar investigations into the solution conformation of several analogues (for example tetragastrin, octagastrin, minigastrin) have been conducted by other groups (77–79, 102). A variety of structural features of these peptides has been characterized, for example, a 3_{10} -helix initiated by a β -turn in des-Trp¹Nle¹²-minigastrin (78).

CONCLUSION

The preceding sections are intended to offer some insights into the NMR methodologies involved in protein structural analysis. Previously structural analysis was in the domain of X-ray crystallography, and although an initial uneasy alliance existed between the two experimental techniques, this has diminished and now X-ray crystallography and NMR are viewed as essentially complementary: one giving a static, the other a dynamic, representation of the protein. Comparisons of three-dimensional protein structures

derived by both techniques have been made, and although there have been instances of significant differences (see the α -bungarotoxin discussion above and 84), generally they are comparable with the minor exception of the orientation of sidechains (for example, 104, 62, 63).

The clear disadvantage of the NMR technique is the small size of the proteins that are amenable to such studies. The recently developed three- and four-dimensional NMR techniques (123, 50, 32), in tandem with the study of proteins into which NMR visible, stable isotopes (^{13}C , ^{15}N , ^2H) have been incorporated (80), have opened up new approaches to studying larger proteins. The resulting increase in the repertoire of accessible NMR experiments imparts the ability to achieve the spectral resolution required to combat the ever increasing complexity of the spectra associated with larger proteins. With these improvements, the probable molecular weight limit for the structure analysis of proteins by NMR spectroscopy is approximately 30 kd. Even with this limitation, it is anticipated that the wealth of structural and dynamic information gleaned by the NMR technique will afford exquisite insights into the biochemistry and biology of these relatively small proteins.

Literature Cited

- Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., et al. 1969. Structure of rhombohedral 2 zinc insulin crystals. *Nature* 224:491-96
- Andersen, W. F., Ohlendorf, D. H., Takeda, Y., Matthews, B. W. 1981. Structure of the cro repressor from the bacteriophage λ and its interaction with DNA. *Nature* 290:754-58
- Arrowsmith, C. H., Pachter, R., Altman, R. B., Iyer, S. B., Jardetzky, O. 1990. Sequence-specific ^1H NMR assignments and secondary structure in solution of *Escherichia coli* trp repressor. *Biochemistry* 29:6332-41
- Arseniev, A. S., Kondakov, V. I., Maiorov, V. N., Bystrov, V. F. 1984. NMR solution spatial structure of short scorpion insectotoxin I_2A . *FEBS Lett.* 165:57-62
- Aue, W. P., Bartholdi, E., Ernst, R. R. 1976. Two-dimensional spectroscopy application to nuclear magnetic resonance. *J. Chem. Phys.* 64:2229-46
- Basus, V. J., Billeter, M., Love, R. A., Stroud, R. M., Kuntz, I. D. 1988. Structural studies of α -bungarotoxin. 1. Sequence-specific ^1H NMR resonance assignments. *Biochemistry* 27:2763-71
- Basus, V. J., Scheek, R. M. 1988. Structural studies of α -bungarotoxin. 2. ^1H NMR assignments via an improved relayed coherence transfer nuclear Overhauser enhancement experiment. *Biochemistry* 27:2772-75
- Bax, A. 1982. *Two-Dimensional Nuclear Magnetic Resonance*, Boston: Reidel
- Bax, A. 1989. Two-dimensional NMR and protein structure. *Annu. Rev. Biochem.* 58:223-56
- Bax, A. 1989. Homonuclear Hartman-Hahn experiments. In *Methods in Enzymology*, ed. N. J. Oppenheimer, T. L. James, 176:159-68. New York: Academic. 530 pp.
- Bax, A., Freeman, R. 1981. Investigation of complex networks of spin-spin coupling by two-dimensional NMR. *J. Magn. Reson.* 44:542-61
- Berg, J. M. 1988. Proposed structure for the zinc binding domains from transcription factor IIIA and related proteins. *Proc. Natl. Acad. Sci. USA* 85:99-102
- Billeter, M., Qian, Y. Q., Otting, G., Muller, M., Gehring, W. J., et al. 1990. Determination of the three-dimensional structure of the *Antennapedia* homeodomain from the *Drosophila* in solution by ^1H nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* 124:183-97
- Bodanszky, A., Ondetti, M. A., Bodanszky, M. 1972. Secretin. V. Sol-

- vent effects and conformational freedom. *J. Am. Chem. Soc.* 94:3600-3
15. Bodanszky, A., Ondetti, M. A., Mutt, V., Bodanszky, M. 1969. Synthesis of secretin. IV. Secondary structure in a miniature protein. *J. Am. Chem. Soc.* 91:944-49
 16. Boelens, R., Ganadu, M. L., Verheyden, P., Kaptein, R. 1990. Two-dimensional NMR studies on des-pentapeptide-insulin. Proton resonance assignments and secondary structure analysis. *Eur. J. Biochem.* 191:147-53
 17. Boelens, R., Scheek, R. M., van Boom, J. H., Kaptein, R. 1987. Complex of lac repressor headpiece with a 14 base-pair lac operator fragment studied by two-dimensional nuclear magnetic resonance. *J. Mol. Biol.* 193:213-16
 18. Boesch, C., Bundi, A., Oppliger, M., Wüthrich, K. 1978. ¹H Nuclear-magnetic-resonance studies of the molecular conformation of monomeric glucagon in aqueous solution. *Eur. J. Biochem.* 91:209-14
 19. Bontems, F., Roumestand, C., Boyot, P., Gilquin, B., Doljansky, Y., et al. 1991. Three-dimensional structure of natural charybdotoxin as an aqueous solution by ¹H-NMR. Charybdotoxin possesses a structural motif found in other scorpion toxins. *Eur. J. Biochem.* 196:19-28
 20. Borgias, B. A., James, T. L. 1989. Two-dimensional nuclear Overhauser effect: complete relaxation matrix analysis. See Ref. 10, pp. 169-83
 21. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., et al. 1988. Monomeric insulins obtained by protein engineering and their medical implications. *Nature* 333:679-82
 22. Breg, J. N., Boelens, R., George, A. V. E., Kaptein, R. 1989. Sequence-specific ¹H NMR assignments and secondary structure of the lac repressor of bacteriophage P22, as determined by two-dimensional ¹H NMR spectroscopy. *Biochemistry* 28:9826-33
 23. Brown, L. R., Büsch, C., Wüthrich, K. 1981. Location and orientation relative to the micelle surface for glucagon in mixed micelles with dodecylphosphocholine. EPR and NMR studies. *Biochem. Biophys. Acta* 642:296-312
 24. Bystrov, V. F., Okhanov, V. V., Miroshnikov, A. I., Ovchinnikov, Y. A. 1980. Solution structure of apamin as derived from NMR study. *FEBS Lett.* 199:113-17
 25. Campbell, I. D., Baron, M., Cooke, R. M., Dudgeon, T. J., Fallon, A., et al. 1990. Structure-function relationships in epidermal growth factor (EGF) and transforming growth factor-alpha (TGF-alpha). *Biochem. Pharmacol.* 40(1):35-40
 26. Carr, M. D., Pastore, A., Gausepohl, H., Frank, R., Roesch, P. 1990. NMR and molecular dynamics studies of the mK12 zinc finger. *Eur. J. Biochem.* 188:455-61
 27. Carver, J. A., Cooke, R. M., Esposito, G., Campbell, I. D., Gregory, H., et al. 1986. A high resolution ¹H NMR study of the solution structure of human epidermal growth factor. *FEBS Lett.* 205(1):77-81
 28. Catterall, W. A. 1988. Structure and function of voltage-sensitive ion channels. *Science* 242:50-61
 29. Churchill, M. E. A., Tullius, T. D., Klug, A. 1990. Mode of interaction of the zinc finger protein TFIIIA with a 5S RNA gene of *Xenopus*. *Proc. Natl. Acad. Sci. USA* 87:5528-32
 30. Clore, G. M., Gronenborn, A. M. 1987. Determination of 3D structures of proteins in solution by nuclear magnetic resonance spectroscopy. *Protein Eng.* 1:275-88
 31. Clore, G. M., Gronenborn, A. M., Brünger, A. T., Karplus, M. 1985. Solution conformations of a heptadecapeptide comprising the DNA binding helix F of the cyclic AMP receptor protein of *Escherichia coli*. Combined use of ¹H nuclear magnetic resonance and restrained molecular dynamics. *J. Mol. Biol.* 186:435-55
 32. Clore, G. M., Wingfield, P. T., Gronenborn, A. M. 1991. High-resolution three-dimensional structure of interleukin 1 β in solution by three- and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry* 30: 2315-23
 33. Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., et al. 1987. The solution structure of human epidermal growth factor. *Nature* 327: 339-41
 34. Demsey, C. 1986. pH Dependence of hydrogen exchange from backbone peptide amides in apamin. *Biochemistry* 25:3904-11
 35. Derome, A. E. 1987. *Modern NMR Techniques for Chemistry Research*, Oxford: Pergamon
 36. Driscoll, P. C., Groeneborn, A. M., Bress, L., Clore, G. M. 1989. Determination of the three-dimensional solution structure of the antihypertensive and antiviral protein BDS-1 from the sea anemone *Anemonia sulcata*: A study using nuclear magnetic resonance and

- hybrid distance geometry-dynamical simulated annealing. *Biochemistry* 28: 2188-98
37. Englander, S. W., Wand, A. J. 1987. Main-chain-directed strategy for the assignment of ^1H NMR spectra of proteins. *Biochemistry* 26:5953-58
 38. Ernst, R. R., Bodenhausen, G., Wokaun, A. 1987. *Principles of Nuclear Magnetic Resonance One and Two Dimensions*. Oxford: Clarendon
 39. Fairall, L., Rhodes, D., Klug, A. 1986. Mapping of the sites of protection on a 5S RNA gene by the *Xenopus* transcription factor IIIA. A model for interaction. *J. Mol. Biol.* 192:577-91
 40. Fogh, R. H., Kem, W. R., Norton, R. S. 1990. Solution structure of neurotoxin I from the sea anemone polypeptide *Stichodactyla helianthus*. A nuclear magnetic resonance, distance geometry, and restrained molecular dynamics study. *J. Biol. Chem.* 265:13016-28
 41. Fogh, R. H., Mabbutt, B. C., Kem, W. R., Norton, R. S. 1989. Sequence specific ^1H NMR assignments and secondary structure in the sea anemone polypeptide *Stichodactyla helianthus* neurotoxin I. *Biochemistry* 28:1826-34
 42. Gadhavi, P. L., Raine, A. R. C., Alefounder, P. R., Laue, E. D. 1990. Complete assignment of the ^1H NMR spectrum and secondary structure of the DNA binding domain of GAL4. *FEBS Lett.* 276:49-53
 43. Gooley, P. R., Blunt, J. W., Norton, R. S. 1984. Conformational heterogeneity in polypeptides cardiac stimulants from sea anemones. *FEBS Lett.* 174:15-19
 44. Gooley, P. R., Blunt, J. W., Norton, R. S. 1988. Effects of pH and temperature on cardioactive polypeptides from sea anemones: A ^1H NMR study. *Biopolymers* 27:1143-57
 45. Gooley, P. R., Caffrey, M. S., Cusanovich, M. A., MacKenzie, N. E. 1990. Assignment of the ^1H and ^{15}N NMR spectra of *Rhodobacter capsulatus* ferrocytochrome c_2 . *Biochemistry* 29: 2778-90
 46. Gooley, P. R., MacKenzie, N. E. 1988. Location of an α -helix in fragment 96-133 from bovine somatotropin by ^1H NMR spectroscopy. *Biochemistry* 27: 4032-40
 47. Gooley, P. R., Norton, R. S. 1985. Specific assignment of resonances in the ^1H nuclear magnetic resonance spectrum of the polypeptide cardiac stimulant anphopleurin-A. *Eur. J. Biochem.* 153: 529-39
 48. Gooley, P. R., Plaisted, S. M., Brems, D. N., MacKenzie, N. E. 1988. De-termination of local conformational stability in fragment 96-133 of bovine growth hormone by high-resolution ^1H NMR spectroscopy. *Biochemistry* 27:802-9
 49. Griesinger, C., Sørensen, O. W., Ernst, R. R. 1985. Two-dimensional correlations of connected NMR transitions. *J. Am. Chem. Soc.* 107:6394-96
 50. Griesinger, C., Sørensen, O. W., Ernst, R. R. 1987. The practical approach to three-dimensional NMR spectroscopy. *J. Magn. Reson.* 73:576-79
 51. Gronenborn, A. M., Bovermann, G., Clore, G. M. 1987. A ^1H -NMR study of the solution conformation of secretin. Resonance assignment and secondary structure. *FEBS Lett.* 215(1):88-94
 52. Haard, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., et al. 1990. Solution structure of the glucocorticoid receptor DNA-binding domain. *Science* 249:157-60
 53. Holak, T. A., Prestegard, J. H., Forman, J. D. 1987. NMR pseudoenergy approach to the solution structure of acyl carrier protein. *Biochemistry* 26:4652-60
 54. Hosur, R. V., Wider, G., Wüthrich, K. 1983. Sequential individual resonance assignments in the ^1H nuclear magnetic resonance spectrum of cardiotoxin V 12 from *Naja mossaibica mossaibica*. *Eur. J. Biochem.* 130:497-508
 55. Hyberts, S. G., Marki, W., Wagner, G. 1987. Stereospecific assignments of side-chain protons and characterization of torsion angles in eglin C. *Eur. J. Biochem.* 164:625-35
 56. Jeener, J. 1971. *Ampere Int. Summer Sch.*, Basko Proje, Yugoslavia
 57. Kaptein, R., Boelens, R., Scheek, R. M., van Gunsteren, W. F. 1988. Protein structure from NMR. *Biochemistry* 2: 5389-95
 58. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., van Gunsteren, W. F. 1985. A protein structure from nuclear magnetic resonance data. Iac repressor headpiece. *J. Mol. Biol.* 182:179-82
 59. Karplus, M. 1963. Vicinal proton coupling in nuclear magnetic resonance. *J. Am. Chem. Soc.* 85:2870-71
 60. Kessler, H., Bermel, W., Müller, A., Pook, K-H. 1985. Modern nuclear magnetic resonance spectroscopy of peptides. In *The Peptides Analysis, Synthesis, Biology*, eds. S. Udenfriend, J. Meienhofer, 7:437-73. New York: Academic
 61. Kessler, H., Gehrke, M., Griesinger, C. 1988. Two-dimensional NMR spectroscopy: Background and overview of the

- experiments. *Angew Chem. Int. Ed. Engl.* 27:490-536
62. Kline, A. D., Braun, W., Wüthrich, K. 1986. Studies by ^1H nuclear magnetic resonance and distance geometry of the solution conformation of the α -amylase inhibitor tendamistat. *J. Mol. Biol.* 189:377-82
 63. Kline, A. D., Braun, W., Wüthrich, K. 1989. Determination of the complete three-dimensional structure of the α -amylase inhibitor tendamistat in aqueous solution by nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* 204:675-724
 64. Kline, A. D., Justice, R. M. 1990. Complete sequence-specific ^1H NMR assignments for human insulin. *Biochemistry* 29:2906-13
 65. Klevit, R. E., Herriott, J. R., Horvath, S. J. 1990. Solution structure of a zinc finger domain of yeast ADRI. *Proteins* 7:215-26
 66. Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., et al. 1989. Solution conformation of conotoxin GI determined by ^1H nuclear magnetic resonance spectroscopy and distance geometry calculations. *Biochemistry* 28:4853-60
 67. Kohda, D., Go, N., Hayashi, K., Inagaki, F. 1988. Tertiary structure of mouse epidermal growth factor determined by two-dimensional ^1H NMR. *J. Biochem.* 103:741-43
 68. Kohda, D., Shimada, I., Miyake, T., Fuwa, T., Inagaki, F. 1989. Polypeptide chain fold of human transforming growth factor- α analogous to those of mouse and human epidermal growth factors as studied by two-dimensional ^1H NMR. *Biochemistry* 28:953-58
 69. Kosen, P. A., Finer-Moore, J., McCarthy, M. P., Basus, V. J. 1988. Structural studies of α -bungarotoxin: Corrections in the primary sequence and X-ray structure and characterization of an isotoxic α -bungarotoxin. *Biochemistry* 27:2775-81
 70. Kuntz, I. D., Thomason, J. F., Oshiro, C. M. 1989. Distance geometry. See Ref. 10, pp. 159-204
 71. Labhardt, A. M., Hunziker-Kuik, E. H., Wüthrich, K. 1988. Secondary structure determination for α -neurotoxin from *Dendroaspis polyepis polyepis* based on sequence-specific ^1H nuclear magnetic resonance assignments. *Eur. J. Biochem.* 177:295-305
 72. Lamerichs, R. M. J. N., Boelens, R., van der Marel, G. A., van Boom, J. H., Kaptein, R., et al. 1989. ^1H NMR study of a complex between the lac repressor headpiece and a 22 base pair symmetric lac operator. *Biochemistry* 28:2985-91
 73. Landschulz, W. H., Johnson, P. F., McKnight, S. L. 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. *Science* 243:1681-88
 74. Lee, M. S., Cavanagh, J., Wright, P. E. 1989. Complete assignment of the ^1H NMR spectrum of a synthetic zinc finger from Xfin. Sequential resonance assignments and secondary structure. *FEBS Lett.* 254:159-64
 75. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., Wright, P. E. 1989. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* 245:635-37
 76. Mabbutt, B. C., Norton, R. S. 1990. Sequential ^1H -NMR assignments and secondary structure of the sea anemone polypeptide anthopleurin-A. *Eur. J. Biochem.* 187:555-63
 77. Mammi, S., Goodman, M., Peggion, E., Foffani, M. T., Moroder, L., et al. 1986. Conformational studies on gastrin related peptides by high resolution ^1H NMR. *Int. J. Peptide Protein Res.* 27:145-52
 78. Mammi, S., Mammi, N. J., Peggion, E. 1988. Conformational studies of human des-Trp $^1\text{Nle}^{12}$ -minigastrin in water-trifluoroethanol mixtures by ^1H NMR and circular dichroism. *Biochemistry* 27(4):1374-79
 79. Mammi, S., Peggion, E. 1990. Conformational studies of human (15-2-aminohexanoic acid) little gastrin in sodium dodecyl sulfate micelles by ^1H NMR. *Biochemistry* 29 (22):5265-69
 80. Markley, J. L. 1989. Two-dimensional nuclear magnetic resonance spectroscopy of proteins: An overview. See Ref. 10, pp. 12-64
 81. McDonald, N., Murray-Rust, J., Blundell, T. 1989. Structure-function relationships of growth factors and their receptors. *Br. Med. Bull.* 45(2):554-69
 82. McGraw, S. E., Craik, D. J., Lindenbaum, S. 1990. Testing of insulin hexamer-stabilizing ligands using theoretical binding, microcalorimetry, and nuclear magnetic resonance (NMR) line broadening techniques. *Pharmaceutical Res.* 7(6):600-5
 83. Nakagawa, S. H., Tager, H. S. 1986. Role of the phenylalanine B25-side chain in directing insulin interaction with its receptor. Steric and conformational effects. *J. Biol. Chem.* 261 (16):7332-41
 84. Netteshim, D. G., Edalji, R. P., Mothson, K. W., Greer, J., Zuiderweg, E. R.

- P. 1988. Secondary structure of complement component C3a anaphylatoxin in solution as determined by NMR spectroscopy: Differences between crystal and solution conformations. *Proc. Natl. Acad. Sci. USA* 85:5036-40
85. Nettesheim, D. G., Klevit, R. E., Drobny, G., Watt, D. D., Krishna, N. R. 1989. Proton nuclear magnetic resonance studies on the variant-3 neurotoxin from *Centruroides sculpturatus* Ewing: Sequential assignment of resonances. *Biochemistry* 28:1548-55
 86. Neuhaus, D., Nakaseko, Y., Nagai, K., Klug, A. 1990. Sequence-specific [¹H] NMR resonance assignments and secondary structure identification for 1- and 2-zinc finger constructs from SW15. *FEBS Lett.* 262:179-84
 87. Noggle, J. H., Schirmer, R. H. 1971. *The Nuclear Overhauser Effect, Chemical Applications*. New York: Academic
 88. Norton, R. S., Cossins, A. I., Kem, W. R. 1989. ¹H NMR study of the solution properties of the polypeptide neurotoxin I from the sea anemone *Stichodactyla helianthus*. *Biochemistry* 28:1820-26
 89. Oas, T. G., McIntosh, L. P., O'Shea, E. K., Dahlquist, F. W., Kim, P. S. 1990. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry* 29:2891-94
 90. Omichinski, J. G., Clore, G. M., Appella, E., Sakaguchi, K., Gronenborn, A. M. 1990. High-resolution three-dimensional structure of a zinc finger from a human enhancer binding protein in solution. *Biochemistry* 29:9324-34
 91. O'Shea, E. K., Rutkowski, R., Kim, P. S. 1989. Evidence that the leucine zipper is a coiled coil. *Science* 243:538-42
 92. O'Shea, E. K., Rutkowski, R., Staford, W. F. III, Kim, P. S. 1989. Preferential heterodimer formation by isolated leucine zippers from *fos* and *jun*. *Science* 245:646-48
 93. Otting, G., Marchot, P., Bougis, P. E., Rochat, H., Wüthrich, K. 1987. Monitoring the purification by high-performance liquid chromatography of cardiotoxins from *Naja mossaibica mossaibica* using phase sensitive two-dimensional nuclear magnetic resonance. *Eur. J. Biochem.* 168:603-7
 94. Otting, G., Steinmetz, W. E., Bougis, P. E., Rochat, H., Wüthrich, K. 1987. Sequence-specific ¹H-NMR assignments and determination of the secondary structure in aqueous solution of the cardiotoxins CTXIIa and CTXIIb from *Naja mossaibica mossaibica*. *Eur. J. Biochem.* 168:609-20
 95. Pabo, C. O., Lewis, M. 1982. The operator-binding domain of λ repressor structure and DNA recognition. *Nature* 298:443-47
 96. Pan, T., Coleman, J. E. 1990. The DNA binding domain of GAL4 forms a binuclear metal ion complex. *Biochemistry* 29:3023-29
 97. Pan, T., Coleman, J. E. 1990. GAL4 transcription factor is not a zinc finger but forms a Zn(II)₂Cys₆ binuclear cluster. *Proc. Natl. Acad. Sci. USA* 87:2077-81
 98. Pan, T., Coleman, J. E. 1989. Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor. *Proc. Natl. Acad. Sci. USA* 86:3145-49
 99. Pardi, A., Galdes, A., Florance, J., Manicote, D. 1989. Solution structures of α -conotoxin GI determined by two-dimensional NMR spectroscopy. *Biochemistry* 28:5494-5501
 100. Pease, J. H. B., Kumar, M. V., Schweitz, H., Kallenbach, N. R., Wemmer, D. E. 1989. NMR Studies of toxin III from the sea anemone *Radianthus paumotensis* and comparison of its secondary structure with related toxins. *Biochemistry* 18:2199-2204
 101. Pease, J. H. B., Wemmer, D. E. 1988. Solution structure of apamin determined by nuclear magnetic resonance and distance geometry. *Biochemistry* 27:8491-98
 102. Peggion, E., Foffani, M. T., Wunsch, E., Moroder, L., Borin, G., et al. 1985. Conformational properties of gastrin fragments of increasing chain length. *Biopolymers* 24:647-66
 103. Pennington, M. W., Kem, W. R., Norton, R. S., Dunn, B. M. 1990. Chemical synthesis of a neurotoxic polypeptide from the sea anemone *Stichodactyla helianthus*. *Int. J. Pept. Protein Res.* 36:335-43
 104. Pflugrath, J. W., Wiegand, G., Huber, R., Vértessy, L. 1986. Crystal structure determination, refinement, and the molecular model of the α -amylase inhibitor Hoe-467A. *J. Mol. Biol.* 189:383-86
 105. Piantini, U., Sørensen, O. W., Ernst, R. R. 1982. Multiple quantum filters for elucidation NMR coupling. *J. Am. Chem. Soc.* 104:6880-81
 106. Povey, J. F., Diakun, G. P., Gramer, C. D., Wilson, S. P., Laue, E. D. 1990. Metal ion coordination in the DNA binding domain of the yeast transcription activator GAL4. *FEBS Lett.* 266:142-46

107. Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J., et al. 1989. The structure of the *Antennapaedia* homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. *Cell* 59: 573-80
108. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., et al. 1983. Improved spectral resolution in COSY ¹H NMR spectra of proteins via double-quantum filtering. *Biochem. Biophys. Res. Commun.* 117(2):479-85
109. Robinson, R. M., Blakeney, E. W. Jr., Mattice, W. L. 1982. Lipid-induced conformational changes in glucagon, secretin, and vasoactive intestinal peptide. *Biopolymers* 21:1217-28
110. Roy, M., Lee, R. W. K., Brange, J., Dunn, M. F. 1990. ¹H NMR spectrum of the native human insulin monomer. Evidence for conformational differences between the monomer and aggregated forms. *J. Biol. Chem.* 265(10):5448-52
111. Saudek, V., Pastore, A., Castiglione Morelli, M. A., Frank, R., Gausepohl, H., et al. 1990. Solution structure of the DNA-binding domain of the yeast transcriptional activator protein GCN4. *Protein Eng.* 4:3-10
112. Schevitz, R. W., Otwinowski, Z., Joachimak, A., Lawson, C. L., Sigler, P. B. 1985. The three dimensional structure of trp repressor. *Nature* 317:782-86
113. Schwabe, J. W. R., Neuhaus, D., Rhodes, D. 1990. Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature* 348:458-61
114. Shaka, A. J., Freeman, R., 1983. Simplification of NMR spectra by filtration through multiple-quantum coherence. *J. Magn. Reson.*, 51:169-73
115. Scheek, R. M., van Gunsteren, W. F., Kaptein, R. 1989. Molecular dynamics simulation techniques for determination of molecular structures from nuclear magnetic resonance data. See Ref. 10, pp. 204-18
116. Steinmetz, W. E., Moonen, C., Kumar, A., Lazdunski, M., Visser, L., et al. 1981. ¹H Nuclear magnetic resonance studies of the conformance of the conference of cardiotoxin V^{H2} from *Naja mossambica mossambica*. *J. Biochem.* 120:467-75
117. Tabate, T., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., et al. 1989. A protein that binds to a *cis*-acting element of wheat histone genes has a leucine zipper motif. *Science* 245:965-71
118. Torda, A. E., Baldwin, G. S., and Norton, R. S. 1985. High-resolution proton nuclear magnetic resonance studies of human gastrin. *Biochemistry* 24(7): 1720-27
119. Torda, A. E., Mabbutt, B. C., van Gunsteren, W. F., Norton, R. S. 1988. Backbone folding of the polypeptide cardiac stimulant anthopleurin-A determined by nuclear magnetic resonance, distance geometry and molecular dynamics. *FEBS Lett.* 239:266-70
120. Torda, A. E., Norton, R. S. 1987. Amide proton exchange rates in radioactive sea anemone polypeptides. *Biochem. Int.* 15:659-66
121. Tracy, H. J., Gregory, R. A. 1964. Physiological properties of a series of synthetic peptides structurally related to gastrin I. *Nature* 204:935-38
122. de Vlieg, J., Scheek, R. M., van Gunsteren, W. F., Berendsen, H. J. C., Kaptein, R., et al. 1988. Combined procedure of distance geometry and restrained dynamics techniques for protein structure determination from nuclear magnetic resonance data: Application to the DNA binding domain of lac repressor from *Escherichia coli*. *Proteins* 3:209-18
123. Vuister, G. W., Boelens, R. 1987. Three-dimensional J-resolved NMR spectroscopy. *J. Magn. Reson.* 73:328-33
124. Wagner, G., Neuhaus, D., Wörgötter, E., Vasak, M., Kägi, J. H. R., et al. 1986. Nuclear magnetic resonance identification of "Half-Turn" and ³10-helix secondary structure in rabbit liver metallothionein-2. *J. Mol. Biol.* 187: 131-35
125. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., et al. 1990. Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. *Nature* 347:575-78
126. Weiss, M. A., Nguyen, D. T., Khait, I., Inouye, K., Frank, B. H., et al. 1989. Two-dimensional NMR and photoCIDNP studies of the insulin monomer: Assignment of aromatic resonances with application to protein folding, structure, and dynamics. *Biochemistry* 28(25): 9855-73
127. Wemmer, D., Kallenbach, N. R. 1983. Structure of apamin in solution: A two-dimensional nuclear magnetic resonance study. *Biochemistry* 22:1901-6
128. Wemmer, D. E., Kumar, N. V., Metrione, R. M., Lazdunski, M., Drobný, G., et al. 1986. NMR analysis and sequence of toxin II from the sea anemone *Radianthus paumotensis*. *Biochemistry* 25:6842-49

129. Wemmer, D. E., Reid, B. R. 1985. High resolution NMR studies of nucleic acids and proteins. *Annu. Rev. Phys. Chem.* 36:105-37
130. Wider, G., Lee, K. H., Wüthrich, K. 1982. Sequential resonance assignments in protein ¹H nuclear magnetic resonance spectra. Glucagon bound to perdeuterated dodecylphosphocholine micelles. *J. Mol. Biol.* 155:367-88
131. Widmer, H., Billeter, M., Wüthrich, K. 1989. Three dimensional structure of the neurotoxin ATX Ia from *Anemonia sulcata* in aqueous solution determined by nuclear magnetic resonance spectroscopy. *Proteins* 6:357-71
132. Widmer, H., Wagner, G., Schweitz, H., Lazdunski, M., Wüthrich, K. 1988. The secondary structure of the toxin ATX Ia from *Anemonia sulcata* in aqueous solution determined on the basis of complete sequence-specific ¹H-NMR assignments. *Eur. J. Biochem.* 171:177-92
133. Wüthrich, K., Billeter, M., Braun, W. 1983. Pseudo-structures for the 20 common amino acids for use in studies of protein conformation by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* 169:949-61
134. Wüthrich, K. 1986. *NMR of Proteins and Nucleic Acids*, New York: Wiley
135. Zagorski, M. G., Bowie, J. U., Verzhom, A. K., Sauer, R. T., Patel, D. J. 1989. NMR studies of arc repressor mutants: protein assignments, secondary structure, and long range contacts for thermostable proline-8 → leucine variant of arc. *Biochemistry* 28:9812-25
136. Zuiderweg, E. R. P., Boelens, R., Kaptein, R. 1985. Stereospecific assignments of ¹H-NMR methyl lines and conformation of valyl residues in the lac repressor headpiece. *Biopolymers* 24: 601-11
137. Zuiderweg, E. R. P., Kaptein, R., Wüthrich, K. 1983. Secondary structure of the lac repressor DNA-binding domain by Two-dimensional ¹H nuclear magnetic resonance in solution. *Proc. Natl. Acad. Sci. USA* 80:5837-41